Niveau(x) d'enseignement: 7 Numéro EDUQ : 8470065
Centre d'intérêt: 570

Allaire, Denis.

Etude de l'effet de méthode entre deux mesures de satisfaction au moyen d'une analyse de variance multitrait-multiméthode / Denis Allaire et Lucie Houde. -- [Montréal] : INRS-Education, [1974].
ii, [31 f. ; 28 cm.

"Communication présentée lors du 42ème Congrès de l'Association canadienne-française pour l'avancement des sciences, Québec, 9 mai 1974". Résumé: f. [1]. Document: 0574-06. Bibliogr.: f. [12-13].

Descripteur(s): *Evaluation de l'enseignant *Evaluation par l'étudiant *Satisfaction de l'étudiant *Relations maître-élèves *Attitude de l'enseignant *Statistique *Test *Collégial (Niveau)

Identificateur(s) de nom: *PERPE (Test)

Identificateur(s) de sujet: *Variance multitrait-multiméthode

---

Contrairement à la majorité des instruments d'évaluation des professeurs, le test PERPE (Perceptions étudiantes de la relation professeur-étudiants) utilise une mesure de satisfaction dérivée. Les étudiants qui évaluent leurs professeurs à l'aide de ce test doivent inscrire deux réponses à chaque item: d'abord le point de l'échelle qui décrit le mieux leur professeur et ensuite le point où ce professeur devrait se situer pour que l'étudiant s'en trouve satisfait. La satisfaction de l'étudiant est mesurée en faisant la différence absolue entre ses réponses aux deux questions précédentes. Le but de la présente étude est de vérifier jusqu'à quel point cette mesure d'écart concorde avec celle obtenue au moyen d'une question portant directement sur la satisfaction de l'étudiant.

Deux formes abrégées du test PERPE ont été construites à partir des mêmes quinze items. L'une utilise la mesure directe de satisfaction et l'autre la mesure dérivée ou d'écart. Durant les mois de mars et avril 1974, l'un et l'autre questionnaires ont été administrés aux étudiants de 51 classes provenant d'un éventail de sept collèges et d'une douzaine de matières différentes. L'analyse des intercorrélations entre les quinze items et les deux méthodes a été réalisée au moyen d'une analyse de variance modifiée. Une discussion est faite de l'application de cette technique à une matrice multitrait-multiméthode, compte tenu de son caractère passablement inusité.

De la variance totale, seulement 4% est attribuable à la différence entre les méthodes, ce qui constitue une claire indication que la mesure dérivée du test PERPE correspond étroitement à une mesure directe de satisfaction. Ces résultats contraires à ceux obtenus par d'autres auteurs sont discutés. Il est suggéré que la convergence entre les deux méthodes, telle qu'observée en utilisant le test PERPE, est peut-être due à certaines directives qui accompagnent ce test.

Type de document: r
Notice(s) connexe(s): 9470096, 9470097

EDUQ 847008

405

DENIS ALLAIRE et LUCIE MONDE, INRS-Éducation, Université du Québec

Etude de l'effet de méthode entre deux mesures de satisfaction au moyen d'une analyse de variance multitrait-multiméthode.

Communication présentée lors du 42ème Congrès de l'Association Canadienne-Française pour l'Avancement des Sciences. Québec, 9 mai 1974.

RESUME

Deux formes abrégées du questionnaire PERPE (Perception par les Etudiants de la Relation Professeur-Etudiants) ont été expérimentées afin de vérifier la convergence entre la mesure dérivée de satisfaction qu'utilise ce test et une mesure directe de satisfaction. L'analyse des intercorrélations entre les 15 items et les deux méthodes a été réalisée au moyen d'une analyse de variance modifiée. Une discussion est faite de l'application de cette technique à une matrice multitrait-multiméthode, compte tenu de son caractère passablement inusité.

De la variance totale, seulement 4% est attribuable à la différence entre méthodes, ce qui constitue une claire indication que la mesure dérivée du test PERPE correspond étroitement à une mesure directe de satisfaction. Ces résultats contraires à ceux obtenus par d'autres auteurs sont discutés. Il est suggéré que la convergence entre les deux méthodes, telle qu'observée en utilisant le test PERPE, est peut-être due à certaines directives qui accompagnent ce test.

(Cette recherche a été rendue possible grâce à une subvention du Ministère de l'Education du Québec, dans le cadre de son "Programme de formation de chercheurs et d'action concertée", accordée au Projet PERPE.)

Le test PERPE (Gagné, 1969; Gagné et Chabot, 1970), contrairement à la majorité des instruments d'évaluation des professeurs, utilise une mesure de satisfaction dérivée. Les étudiants qui évaluent leurs professeurs à l'aide de ce test doivent inscrire deux réponses à chaque item: d'abord le point de l'échelle qui décrit le mieux leur professeur et ensuite le point où ce professeur devrait se situer pour que l'étudiant s'en trouve satisfait. Un item-type de PERPE est illustré à la figure 1. La satisfaction de l'étudiant est mesurée en faisant la différence absolue entre ses réponses aux deux questions précédentes.

Le but de la présente étude est de vérifier jusqu'à quel point cette mesure d'écart concorde avec celle obtenue au moyen d'une question portant directement sur la satisfaction de l'étudiant.

DESCRIPTION DE L'EXPÉRIENCE

Deux formes abrégées du test PERPE, présentées à l'Appendice A, ont été construites à partir des mêmes 15 items. L'une utilise la mesure

32. HUMOUR: l'humour en classe peut servir tout aussi bien à détendre l'atmosphère qu'à aider le professeur à capter l'attention du groupe au plan du contenu: il devient alors "outil pédagogique". Cet humour peut prendre la forme de reparties, facéties, jeux de mots, calembours, etc.

1. AUCUN humour
2. Humour TRÈS RARE
3. Humour OCCASIONNEL
4. Humour FRÉQUENT
5. Humour TRÈS FRÉQUENT

Question A: Où situez-vous ce cours (ce professeur) sur l'échelle d'évaluation?

Question B: Où devrait-il se situer pour que vous soyez satisfait(e)?

Figure 1. Item-type du test PERPE (version 1970 à 1974).

directe de satisfaction et l'autre la mesure dérivée ou d'écart. Durant les mois de mars et avril 1974, l'un et l'autre questionnaire ont été administrés aux étudiants de 51 classes provenant d'un éventail de sept collèges et d'une douzaine de matières différentes.

ANALYSE ET RESULTATS

Pour chaque professeur de l'échantillon, la moyenne des réponses d'insatisfaction de ses étudiants a été obtenue pour chacun des 15 traits (ou items) selon chacune des deux méthodes.

Les intercorrélations entre ces 30 mesures constituent une matrice multitrait-multiméthode semblable à celle utilisée par Campbell et Fiske (1959) pour vérifier les critères de validité convergente (le fait que des méthodes différentes mesurent la même chose) et discriminante (le fait que des traits différents mesurent des choses différentes). Selon ces auteurs, il y aura validité convergente et discréminante à la condition que les corrélations monotrait-hétérométhodes soient supérieures aux corrélations hétérotrait-monométhodes et hétérotrait-hétérométhodes. En guise de test de validation, ils suggèrent donc de vérifier si les intercorrélations de la matrice se conforment de façon générale à ces exigences.

Ce genre d'analyse correspond beaucoup plus à un examen ou à une inspection de la matrice multitrait-multiméthode, qu'à son analyse statistique au sens strict. En effet, l'examen informel qu'ils font des intercorrélations ne rencontre aucun des deux buts habituels d'une véritable analyse statistique, qui sont, premièrement, de tester une hypothèse et, deuxièmement, de produire les statistiques descriptives capables de résumer les données à leur plus simple expression.

Pour pallier aux insuffisances de l'approche de Campbell et Fiske, une analyse[1] de variance "personnes x traits x méthodes" a été réalisée

---

[1] Outre l'analyse de variance, l'analyse factorielle a également été proposée pour l'analyse d'une matrice multitrait-multiméthode. Brunet (1974) résume les travaux portant sur l'application de cette seconde technique.

sur les données de cette expérience. Au niveau inférentiel, il sera possible de vérifier si les deux méthodes sont significativement différentes en testant l'interaction "personnes x méthodes", tandis qu'au niveau descriptif, la part de variance associée à cette source pourra être établie.

Une utilisation passablement similaire de l'analyse de variance avait déjà été proposée par Guilford (1954) et Willingham et Jones (1958), dans une situation multitrait-multijuge, afin de vérifier au moyen de l'interaction "personnes x juges" la présence d'un effet de halo. Stanley (1961), Cronbach et al. (1972) et Allaire (1973a) ont par la suite suggéré d'utiliser l'analyse de variance dans le cas d'une situation multitrait-multiméthode. Ces derniers auteurs ont cependant oublié de discuter un des problèmes majeurs que préserte l'analyse de variance dans ce genre de situation expérimentale. Contrairement aux mesures issues de plusieurs juges, celles issues de plusieurs méthodes utilisent habituellement des échelles de mesures différentes, ce qui constitue un viol évident du postulat d'homogénéité des variances. Il est clair que s'il existe une différence de variance entre les scores obtenus par deux méthodes, ce seul fait suffit à produire une interaction significative, alors même que les deux méthodes seraient en corrélation parfaite.

Pour contrevenir à cette difficulté, les scores des professeurs à chacune des 30 mesures ont été ramenés sur une même échelle (de moyenne égale à 0.0 et de variance égale à 1.0). Cette standardisation préalable de scores, de façon à rendre possible l'analyse de variance d'un plan comportant plusieurs traits et plusieurs méthodes, n'a été suggérée qu'à quelques rares reprises. Zyzansky (1962), Boruch et al. (1970) et Kavanagh et al. (1971) sont les seuls, à notre connaissance, à l'avoir utilisée. La seule différence entre l'analyse réalisée ici et les leurs est que nous avons considéré les facteurs "personnes", "traits" et "méthodes" comme aléatoires, alors que ces auteurs ont conçu les traits et les méthodes comme fixes.

McNemar (1958) utilise également l'analyse de variance, après standardisation des mesures, pour une situation d'expérience se rapprochant de la nôtre; soit celle où deux formes parallèles sont utilisées selon deux

méthodes. Outre ces auteurs, la littérature ne semble guère contenir d'autres exemples de ce genre d'analyse, et encore moins de critiques à son emploi. Il n'y aurait donc d'autres conditions à son utilisation que celles habituelles à l'analyse de variance. Ainsi, comme McNemar (1958) le mentionne, le postulat de l'homogénéité des variances d'erreur exige ici que les fidélités de chaque trait et méthode soient égales (parce que $\sigma_e^2 = 1 - \rho_{xx'}$). Les résultats de fidélité rapportés par Allaire (1973b), sur des questionnaires presque identiques à ceux de la présente expérience, indiquent que les deux méthodes ne diffèrent guère quant à la fidélité.

Il est à noter que la transformation des scores qui a été réalisée ici a pour effet de simplifier considérablement les calculs de l'analyse de variance. Premièrement, le fait de ramener à une moyenne de zéro les différents traits et méthodes amène la disparition des carrés moyens "traits", "méthodes" et "traits x méthodes". En second lieu, les quatre carrés moyens restant peuvent directement être calculés à partir des moyennes respectives des corrélations (1) hétérotrait-monométhodes, (2) monotrait-hétérométhodes et (3) hétérotrait-hétérométhodes. Cela tient au fait que ces carrés moyens peuvent s'exprimer en termes de covariances (voir Stanley, 1961, formules 1 à 4), lesquelles correspondent à des corrélations une fois les mesures réduites à une variance de 1.0.

La figure 2 montre les moyennes des divers types d'intercorrélations de la matrice multitrait-multiméthode ayant servi au calcul[1] de l'analyse de variance présenté au tableau 1. L'examen de ces résultats montre une interaction "professeurs x méthodes" significative ($p < .01$). Les deux méthodes ne sont donc pas parfaitement identiques. Cependant, un effet qui s'avère significatif n'implique pas nécessairement qu'il s'agisse d'un effet important, surtout quand le nombre de degré de liberté associés au test est grand, comme c'est ici le cas.

---

[1] Les formules nécessaires à l'analyse de variance d'une matrice multitrait-multiméthode sont présentées à l'Appendice B.

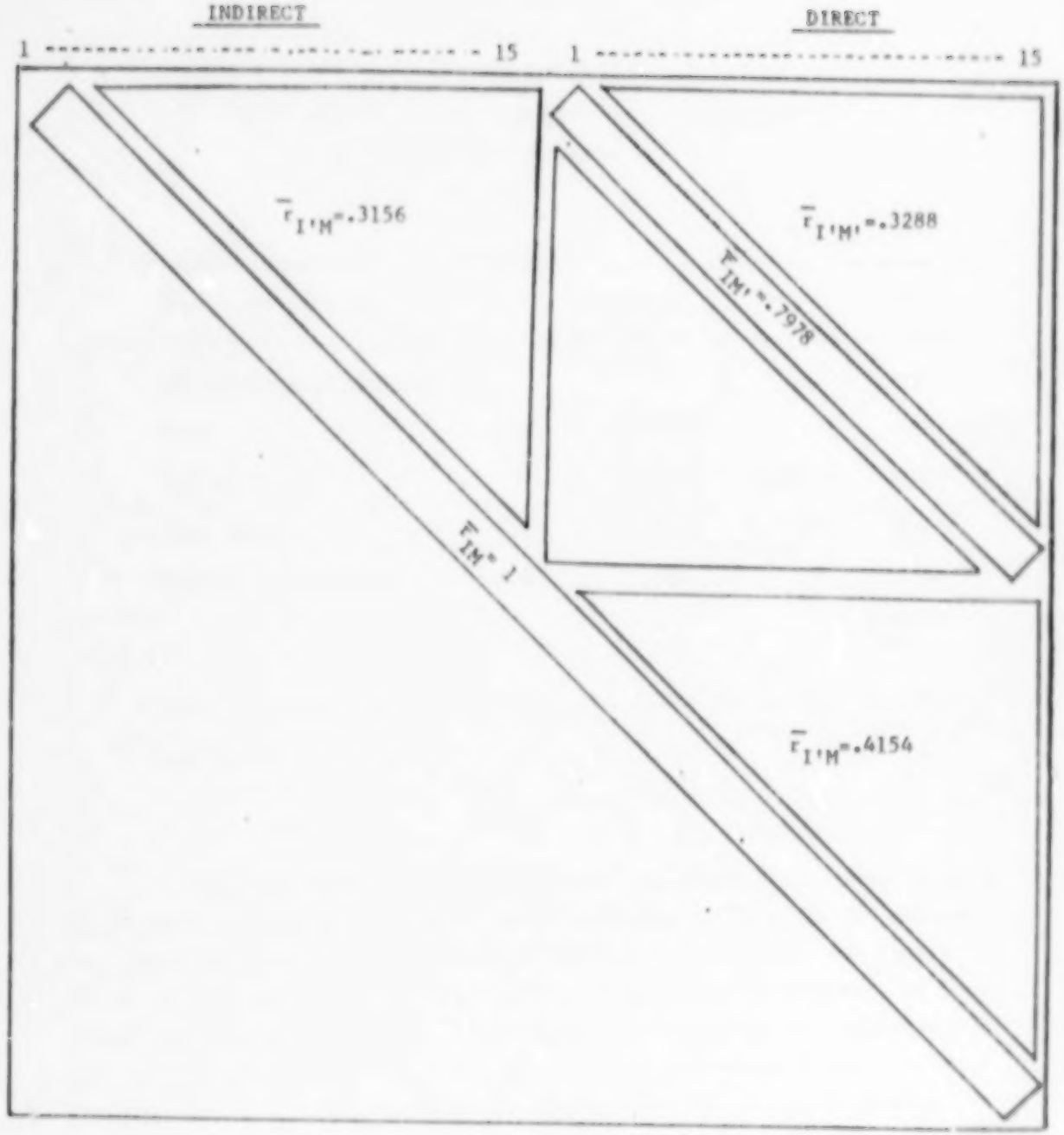

Figure 2. Moyennes des intercorrélations de la matrice multitrait-multiméthode.

Tableau 1

ANOVA multitrait-multiméthode sur les données recueillies au printemps 1974

Source de variance	d.l.[a]	C.M.	F	$\hat{\sigma}^2$
Satisfaction vs Ecart				
Prof.	50	11.5180		.3287
Prof.x Items	700	1.1035	6.668**	.4690
Prof.xMéth.	50	.7160	4.326**	.0367
Prof.x Items x Méth.	700	.1655		.1655
				1.0

[a] $n_p = 51$, $n_i = 15$, $n_m = 2$

Plus instructives à cet égard sont les composantes de variance de la dernière colonne du tableau 1. De la variance totale, 33% est attribuable aux différences entre professeurs, 47% à l'hétérogénéité des traits, 17% à l'erreur de mesure (ou à la triple interaction) et seulement 4% à l'hétérogénéité des méthodes. C'est dire que l'influence des méthodes est ici très restreinte. Si cette variance est supérieure à zéro, elle demeure cependant très inférieure à celle des traits qu'il y aura plus de ressemblance entre les mesures du même trait par deux méthodes qu'entre les mesures de deux traits par la même méthode. Cette dernière inégalité apparaît donc confirmer la validité de ces mesures, du moins en ce qui a trait à certaines des conditions formulées par Campbell et Fiske (i.e. les deuxième et troisième conditions exprimées aux pp. 82-83 de leur étude).

DISCUSSION

Wanous et Lawler (1972) ont réalisé une étude multitrait-multiméthode afin de valider diverses mesures de satisfaction face à divers aspects d'un emploi. Deux de leurs mesures sont identiques à celles expérimentées ici: soit une mesure directe et une mesure dérivée ("Should be - Is now"). Nous avons extrait de leur matrice multitrait-multiméthode les intercorrélations relatives à ces deux méthodes de mesure et avons réalisé la même analyse de variance que précédemment. Une seconde analyse de variance a également été réalisée en substituant à la mesure dérivée la mesure de "réalité" (Is now). Les résultats de l'une et l'autre analyse apparaissent au tableau 2.

Une première différence entre leurs résultats et les nôtres, est l'importance plus grande de la variance résiduelle (47% et 35% contre 17%). Cependant, elle s'explique aisément par le fait que les mesures de Wanous et Lawler seraient moins fidèles que les nôtres, compte tenu qu'elles originent non pas des moyennes de plusieurs réponses mais de réponses individuelles.

Une seconde différence, du moins dans le cas de la mesure d'écart, est l'importance beaucoup plus grande des méthodes. Alors que seulement 4% de la variance non-résiduelle (.0367/(1-.1655)) est attribuable à cette source dans nos résultats, ce pourcentage est de 29 (.1544/(1-.4739)) dans les leurs, ce qui indique une faible convergence entre leurs mesures de satisfaction directe et de satisfaction dérivée. Par contre, leur mesure de réalité apparaît beaucoup plus liée à celle de satisfaction directe. Seulement 7% de la variance non-résiduelle est attribuable aux méthodes dans la seconde analyse de variance où ces deux mesures sont comparées.

Leurs résultats, contrairement aux nôtres, semblent donc remettre en cause l'équivalence entre mesures directe et dérivée de satisfaction. Ces auteurs ne sont pas d'ailleurs les seuls à présenter des résultats en ce sens. Imparato (1972) démontre qu'une même valeur d'écart (entre la réalité et l'idéal) produit des degrés de satisfaction variables selon que l'idéal est élevé ou pas, alors qu'il ne réussit pas à montrer que la

Tableau 2

ANOVA multitrait-multiméthode d'après les corrélations du tableau 3 de Wanous et Lawler (1972).

Source de variance	d.l.[a]	C.M.	F	δ^2
<u>Satisfaction vs Ecart</u>				
Pers.	207	3.8267		.1683
Pers. x Items	1035	.8807	1.858**	.2034
Pers. x Méth.	207	1.4003	2.955**	.1544
Pers. x Items x Méth.	1035	.4739		.4739
				1.0
<u>Satisfaction vs Réalité</u>				
Pers.	207	4.6615		.2833
Pers. x Items	1035	.9997	2.887**	.3267
Pers. x Méth.	207	.6085	1.757**	.0437
Pers. x Items x Méth.	1035	.3463		.3463
				1.0

** prob. $<.01$

[a] $n_p = 208$, $n_i = 6$, $n_m = 2$

satisfaction d'un répondant est affectée par l'écart entre ses réponses de réalité et d'idéal. Hall (1970) montre que les mesures de réalité prédisent mieux la satisfaction d'étudiants face à ce qu'ils ont appris que les mesures dérivées de satisfaction. Levinthal (1974) n'obtient de corrélations significatives entre mesures directe et dérivée que pour une minorité des items d'un questionnaire d'évaluation de l'enseignement.

La raison à cette disparité entre les résultats précités et les nôtres réside peut-être dans les directives qui accompagnent le test PERPE. En effet, les répondants à ce test sont avertis que tout écart entre leurs deux réponses sera interprété comme de l'insatisfaction, et toute coincidence entre ces mêmes réponses comme de la satisfaction. Il serait donc étonnant, dans le cas du test PERPE, compte tenu de ces consignes explicites, qu'un minimum de convergence entre les deux méthodes ne puisse être observé. La question demeure cependant de savoir à quel point la mesure dérivée, sans le secours de semblables directives, présentera encore la même convergence avec une mesure directe de satisfaction. Les résultats rapportés par les auteurs précédents suggèrent qu'il peut en être autrement.

APPENDICE B

I : Formules des carrés moyens en termes de corrélations

$$CMp = 1 + (n_i - 1)\,\bar{r}_{i'm} + (n_m - 1)\,\bar{r}_{im'} + (n_i - 1)(n_m - 1)\,\bar{r}_{i'm'}$$

$$CMpi = 1 - \bar{r}_{i'm} + (n_m - 1)\,\bar{r}_{im'} - (n_m - 1)\,\bar{r}_{i'm'}$$

$$CMpm = 1 + (n_i - 1)\,\bar{r}_{i'm} - \bar{r}_{im'} - (n_i - 1)\,\bar{r}_{i'm'}$$

$$CMpim = 1 - \bar{r}_{i'm} - \bar{r}_{im'} - \bar{r}_{i'm'}$$

où $\bar{r}_{i'm}$: la moyenne des corrélations hétérotrait-monométhodes

$\bar{r}_{im'}$: la moyenne des corrélations monotrait-hétérométhodes

$\bar{r}_{i'm'}$: la moyenne des corrélations hétérotrait-hétérométhodes

II : Espérances mathématiques des CM (modèle aléatoire)

$$E(CMp) = \sigma^2(e) + \sigma^2(pim) + n_m\sigma^2(pi) + n_i\sigma^2(pm) + n_m n_i\sigma^2(p)$$

$$E(CMpi) = \sigma^2(e) + \sigma^2(pim) + n_m\sigma^2(pi)$$

$$E(CMpm) = \sigma^2(e) + \sigma^2(pim) + n_i\sigma^2(pm)$$

$$E(CMpim) = \sigma^2(e) + \sigma^2(pim)$$

III: Estimés des composantes de variance

$$\hat{\sigma}^2(p) = \bar{r}_{i'm'}$$

$$\hat{\sigma}^2(pi) = \bar{r}_{im'} - \bar{r}_{i'm'}$$

$$\hat{\sigma}^2(pm) = \bar{r}_{i'm} - \bar{r}_{i'm'}$$

$$\hat{\sigma}^2(pim,e) = CMpim$$

Références

ALLAIRE, D. (1973) Plan factoriel i x j:(m x p) pouvant servir à valider la mesure indirecte de satisfaction du test PERPE. Document 0573-04, INRS-Education, Montréal (a).

ALLAIRE, D. (1973) La fidélité d'une mesure directe et d'une mesure indirecte de l'insatisfaction. Thèse de M.A., Département de Psychologie, Université du Québec à Montréal (b).

BRUNET, R. (1974) Validité convergente et validité discriminante: revue de la méthodologie. Document 0574-05, INRS-Education, Montréal.

CAMPBELL, D.T. & FISKE, D.W. (1959) Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 56, 81-105.

CRONBACH, L.J., GLASER, G.C., NANDA, H. & RAJARATNAM, N. (1972) The dependability of behavioral measurements. New York: Wiley.

GAGNE, F. (1969) Qu'est-ce que le test PERPE? Document 1169-02. Projet PERPE.

GAGNE, F. et CHABOT, M. (1970) PERPE: une conception neuve du perfectionnement pédagogique. Prospectives, 6(13), 160-181.

GUILFORD, J.P. (1954) Psychometric methods. New York: McGraw-Hill.

HALL, D.T. (1970) The effect of teacher-student congruence upon student leraning in college classes. Journal of Educational Psychology, 61(3), 205-213.

IMPARATO, N. (1972) Relationship between Porter's Need Satisfaction Questionnaire and the Job Descriptive Index. Journal of Applied Psychology, 56(5), 397-405.

LEVINTHAL, C.F. (1974) An analysis of the teacher evaluation process. Final report, project no. 2B089, U.S. Department of Health, Education and Welfare. National Institutes of Education, Washington, D.C.

MCNEMAR, Q. (1958) Attenuation and interaction. Psychometrika, 23, 259-266.

STANLEY, J.C. (1961) Analysis of unreplicated three-way classifications, with applications to rater bias and trait independance. Psychometrika, 26(2), 205-219.

WANOUS, J.P. & LAWLER, E.E. (1972) Measurement and meaning of job satisfaction. Journal of Applied Psychology, 56(2), 95-105.

WILLINGHAM, W.W. & JONES, M.B. (1958) On the identification of halo through analysis of variance. Educational and Psychological Measurement, 18(2), 403-407.

ZYZANSKI, S.J. (1962) Analysis of variance applied to factors which do not have comparable scales. M.S. Thesis, Iowa State University of Science and Technology. Ames: Iowa.